Mass-deployable Smartphone-based Objective Hearing Screening with Otoacoustic Emissions

Nils Heitmann
Technical University of Munich
Munich, Germany
heitmann@rcs.ei.tum.de

Thomas Rosner
PATH medical GmbH
Germering, Germany
rosner@pathme.de

Samarjit Chakraborty
University of North Carolina at
Chapel Hill (UNC), USA
samarjit@cs.unc.edu

ABSTRACT

Since hearing loss is one of the most widespread disabilities and can often be addressed by early detection and intervention, there is a strong interest in technologies for cost-effective and mass-deployable hearing *screening*. Towards this, smartphones have been used for *subjective* tests where a sequence of tones are played to a subject who has to appropriately respond upon hearing them. But such tests are inappropriate where, *e.g.*, children are involved who cannot provide reliable feedback, or the test takes too long. In this paper, we investigate an alternative modality to develop an *objective* screening test using smartphones. It relies on how the cochlea actively distorts tones emitted into the ear. By measuring these distorted signals, it is possible to reliably deduce the subject's hearing health. But smartphones are not designed to detect such low signals, and the suitability of a phone depends on the signal processing characteristics of the phone's hardware. In this paper we investigate this issue in detail and conclude that some smartphones are suitable for *objective* screening tests that require no interaction with a subject. This opens up new screening options that were not available before and have immense societal implications in developing countries.

CCS CONCEPTS

• **Applied computing** → **Life and medical sciences**; • **Human-centered computing** → *Interaction techniques.*

KEYWORDS

smartphones, otoacoustic emissions, hearing loss

ACM Reference Format:
Nils Heitmann, Thomas Rosner, and Samarjit Chakraborty. 2021. Mass-deployable Smartphone-based Objective Hearing Screening with Otoacoustic Emissions. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21), October 18–22, 2021, Montréal, QC, Canada.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3462244.3479920

1 INTRODUCTION

With 430 million people suffering from moderate to high hearing loss, it is one of the most widespread disabilities in the world [15]. Even a moderate loss can impact communication, well-being, quality of life, and health. This problem is more severe in developing and under-developed countries. In a study among Indian school children [12] aged between 12-14 years, researchers screened 1,030 urban children and found that 6% of them suffered from some kind of hearing loss compared to nearly 33% in the rural group of 640 children. The difference might be attributed to the lower socio-economic status of the rural population often leading to malnutrition, poorer health education and inadequate medical facilities which all increase the risk of hearing problems. Also, according to a study carried out by the Society to Aid the Hearing Impaired, in many cities in India (such as Hyderabad and Kolkata) 3 out of 4 traffic patrol officers suffered from some form of hearing loss (sometimes even permanent ones). Similar numbers have been found in industrial environments where often safety gears are not used or norms are not strictly followed.

In all of these cases, the deterioration is *progressive* and timely detection and intervention can reduce or completely address the problem. However, in such scenarios there is also a lack of suitable medical facilities and personnel, and cost is an added constraint. To address these, the goal of our research has been to develop a (i) low-cost *hearing screening* device, that (ii) can be operated by a lay person without any medical training. We envision such a device to be used in rural schools, where the class teacher might screen every child once a year, and refer to a doctor for more detailed examination of the test fails. Or they could be used by construction workers or traffic police and detect the onset of hearing issues in a timely manner.

In this paper, we ask *whether a smartphone could be a suitable platform for such a hearing screening device*? Today, smartphones have very high penetration in both developing and under-developed countries and would address requirement (i). Not surprisingly, there are already multiple smartphone apps for hearing screening; see [1] for a review. However, all of these apps rely on what is referred to as *subjective* tests, where the smartphone plays a sequence tones through a headphone into the subject's ear, who has to give feedback on hearing the tone. However, such tests are not suitable, for example, in a rural school for various reasons – children cannot give reliable feedback, each test takes too long, and there is often considerable ambient noise. Hence, we ask a more specific question in this paper: Whether a smartphone could be a suitable platform for *objective* screening tests?

Technical challenges: Unlike subjective tests, objective ones do not require any interaction with the subject and are much faster. While more details are in the subsequent sections, here we are concerned with objective tests based on otoacoustic emissions (OAEs). Here, acoustic stimuli are emitted into the subject's ear canal and

ICMI '21, October 18–22, 2021, Montréal, QC, Canada

ACM ISBN 978-1-4503-8481-0/21/10…$15.00
https://doi.org/10.1145/3462244.3479920

based on how these tones are distorted by the active processes in the cochlea (in the inner ear), a characteristic acoustic emission can be detected in the ear canal from which the hearing (dis)ability of the subject can be deduced. Although there are commercially-available medical devices based on this principle, they are expensive (in the range of five to ten thousand dollars and upwards) and can only be used by trained medical personnel. Our goal is to investigate whether a smartphone with a simple interface could instead provide reliable results. This is fraught with some technical challenges – unlike smartphone-based subjective tests, we now need the input (microphone) path of the smartphone to detect a signal that is well below normal ambient noise. Since the audio subsystems of smartphones are not designed for detecting such low signals, it is not clear whether a smartphone can at all be used, and what kind of signal processing techniques could be necessary for detecting such signals. Second, we aimed to use the standard Android application programming interface (API) in order to develop a general enough solution. Unfortunately, it offers little access to the smartphone's hardware, making our problem more difficult.

Contributions and outcomes: We evaluated seven off-the-shelf smartphones and found that four of them are suitable for OAE-based hearing screening. While all the smartphone could generate suitable stimulus signals, their return paths had very different characteristics. First, we had to check the sound pressure in the ear canal (at the microphone) and what digital values it resulted from the analog-to-digital converter (ADC) in the smartphone. This input-signal path behaved in a non-linear fashion in some smartphones (Motorola G4 Play and Fairphone FP 3), rendering them unsuitable. Second, the reflected signal from the cochlea has to be filtered from ambient noise. This can be done using standard methods (such as synchronous averaging) as long as the phone does not introduce additional distortions. Because OAEs themselves are also distortions, in the presence of additional distortions introduced by the smartphone, we will not be able to reliably detect them. One of our evaluated smartphones (Huawei Nova) failed this requirement. Among the evaluated phones, the Samsung Tab A 10.1, LG G5 SE, Sony Xperia Z3+ and Google Nexus 7 passed all requirements and turned out to be suitable.

Using suitable signal processing algorithms and by comparing with results from standard medical devices, we conclude that smartphone-based *objective* hearing screening is feasible. Our interface – with a simple "pass"/"refer" outcome – is suitable for use by a lay person with little to no training. Hence, we believe that this could result in a solution for mass-deployable hearing screening in developing countries, where a low-cost hearing screening solution, without involving trained medical personnel, is necessary.

To the best of our knowledge, this is the first paper that investigates smartphone-based *objective* hearing screening. Combined with existing subjective screening apps that use different interaction modalities with a subject, *viz.,* explicit feedback, this work opens up more possibilities. Our proposed objective test can be administered very rapidly. If it fails, then a more detailed and carefully orchestrated subjective test, again using a smartphone, but with a different app and headset, can be administered. Subsequently, the subject might be referred to a medical practitioner if needed. But a first rapid objective test significantly enlarges the pool of subjects who can now be tested. As explained in the paper, there are also opportunities for improving the signal processing, to further reduce the screening time per subject and bringing the performance closer to that of specialized medical-grade screening devices.

In the next section we introduce the relevant concepts of hearing screening and explain the used OAE measurement protocols. Section 3 discusses how we set up the different smartphones in order to evaluate them for objective hearing screening. In Section 4 the audio input and output paths of the smartphones are characterized. Finally, in Section 5 we conduct OAE measurements with the smartphones and compare the results with those obtained using a medical-grade equipment.

2 BASICS OF HEARING SCREENING

In audiology, hearing tests can be categorized into *diagnostic* and *screening* tests. Diagnostic tests are used in a clinical setting to identify diseases and find appropriate treatments. Screening tests, however, are intended to quickly evaluate the state of the auditory system, to discern if a diagnostic test is necessary. The result of such a test is either a "*pass*" if the hearing is normal, or a "*refer*" if a normal hearing could not be established. In the latter case, the patient is referred to a doctor for a detailed diagnostic test, since the cause of a "refer" could also be, for example, a faulty test, a misapplied test, or adverse test conditions (*e.g.*, high ambient noise levels).

Active participation in a hearing screening test is not always wanted or is even possible, *e.g.* when testing small children, or in crowded public spaces such as in a school. As outlined in Section 1, audiological methods can be classified as either *subjective* or as *objective* tests. In a subjective hearing test, the patient gives feedback, *e.g.*, by pressing a button if the tone can be heard. For subjective tests, multiple clinically validated smartphone apps are available [1]. Calibrated noise makers, which are intended to be used by medical professionals, have also been proposed as a low-cost solution [16].

Objective hearing tests do not require any feedback from the patient. Common methods are auditory brainstem response (ABR), where neural activity is measured with electrodes on the scalp in response to an acoustic stimulus. A low-cost ABR has also been proposed [19]; however, handling the electrodes and the longer measurement time is often prohibitive in first-level hearing screening procedures, especially in settings we are interested in. Another method involves measuring OAEs, which, in contrast to ABR, only requires placing a single acoustic probe in the ear canal. If normal hearing is present, an OAE test will only take a few seconds in most cases. While measuring OAEs has its caveats, it is often recommended in many hearing screening applications [14] and is particularly suitable in our setups, barring the high cost of OAE equipment and the need for trained medical personnel for operating them and administering the test.

2.1 Otoacoustic Emissions

OAEs are acoustic emissions, which are a result of active amplification processes in the cochlea located in the inner ear. These acoustic emissions travel backwards through the auditory system and can be observed as a minuscule acoustic signal in the ear canal. Since the the mid 70s, when David Kemp [7] was able to first measure this phenomenon, OAEs became common in audiological practice. At

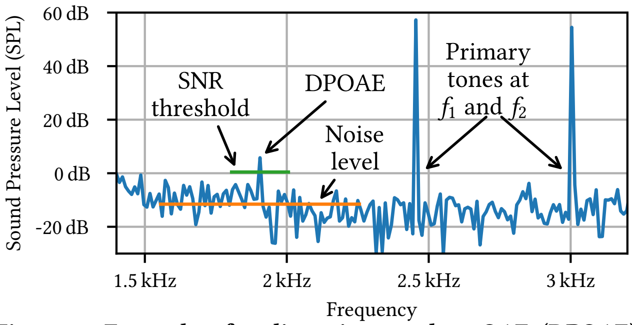


Figure 1: Example of a distortion product OAE (DPOAE) recording at 3 kHz with the Samsung Tab A 10.1. The OAE response has an amplitude of 6 dB SPL and the noise floor is at -12 dB SPL resulting in an SNR of 18 dB.

the same time, advances in technology have made the measurement of OAEs more practical. Today, commercially-available test equipment is often offered as dedicated handheld devices. Connected to such a device is an ear probe, which contains one or more speakers to generate a stimulus to evoke the OAEs, and also a microphone to record the resulting signal. The shape of the ear probe is similar to in-ear headphones. When measuring OAEs, the ear probe is fitted into the ear canal with a replaceable tip, to create a tight seal from the outside environment. This is needed to reduce ambient noise from disturbing the measurement, and also to keep the signal energy of the minuscule OAE signals inside the ear canal. The probe design and fit is crucial for measurement success. When placing the OAE ear probe in the patient's ear, attention must be paid to ensure a proper probe fit. As a part of this research, we also plan to address this issue by designing ear probes equipped with sensors to provide live feedback on the correctness of the fit. But this problem is not the focus of this paper.

Measurement protocols for OAE hearing screening are well established and in this paper we will focus on using these existing screening protocols on a *new platform*, *viz.,* an off-the-shelf smartphone instead of dedicated electronics platform. As discussed in Section 1, our goal is to enable smartphone based objective hearing screening to increase the accessibility by lowering the cost and improving the usability by laypersons. The two most common OAE measurement protocols are DPOAE and TEOAE. We will briefly introduce both protocols, explaining how they are used for calculating the screening results. We rely on the details of these protocols in Section 5 that outlines our findings.

Distortion Product OAEs. A DPOAE response is evoked by stimulating the inner ear with two (primary) pure tones f_1 and f_2. If the cochlea is healthy, a characteristic tone at $2f_1 - f_2$ will be generated by it. Figure 1 shows an example recording in the frequency domain after a fast Fourier transform (FFT). To decide whether the recorded signal at $2f_1 - f_2$ is an actual OAE response or is noise, the noise level needs to be estimated by averaging the surrounding frequency bins. The ratio of $2f_1 - f_2$ and the estimated noise level is the signal to noise ratio (SNR). Only a SNR above a certain threshold, will result in a "pass".

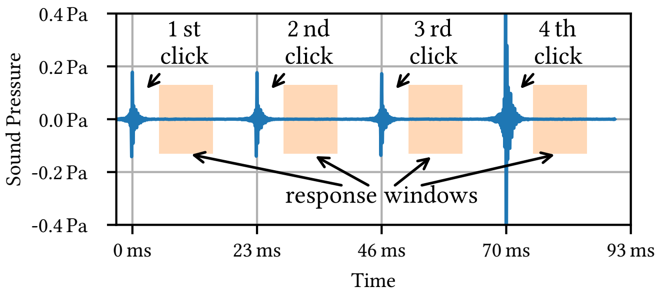
One DPOAE measurement will only test the cochlea at one specific frequency (f_2). Therefore, a small series of measurements at different frequencies must be conducted for a full screening test. The remaining base parameters for an individual measurement are: the frequency of f_1, which is usually defined close to the ratio $f_2/f_1 = 1.22$, and the level (sound pressure) of the primary tones which is set according to the *scissor paradigm* [10].

Transient-evoked OAEs. TEOAEs are excited by a stimulus consisting of a click of typically 100 µs duration and 80 dB peSPL (peak-equivalent sound pressure level [11]) in amplitude. The response evoked in the inner ear by this broadband signal will be detectable in the ear canal after a few milliseconds. To distinguish the non-linear OAEs from the linear components and the echo of the stimulus, a non-linear measurement protocol with a series of clicks is used [8] as follows. Three clicks with normal amplitude are followed by a forth click with three times the amplitude and inverted polarity. Figure 2(a) shows a recorded response of this click sequence. By summing the responses of the four clicks, the linear components are canceled out and only the non-linear OAE, as well as the random noise, remain. Analogous to measuring DPOAE, the noise floor and signal amplitude need to be quantified. Multiple responses are captured, which are then compared to each other. The signal component is approximated by taking the average of all responses, thus lowering the random noise. The noise is estimated by calculating a standard deviation value for each sample in time across all captured responses. Figure 2(b) shows the result of a normal hearing ear after 200 repeated click sequences. The OAE will arrive at the ear probe right after stimulus onset and is evaluated after the stimulus artifact has decayed. By taking the root mean square (RMS) of the extracted signal and the noise inside this window, we obtain single values for signal and noise levels. The SNR can now be calculated and compared to a threshold, in the same manner as with the DPOAE measurement to obtain a "pass" or a "refer". Due to the broadband nature of the stimulus, only one measurement is needed to test the cochlea for a wide range of frequencies.

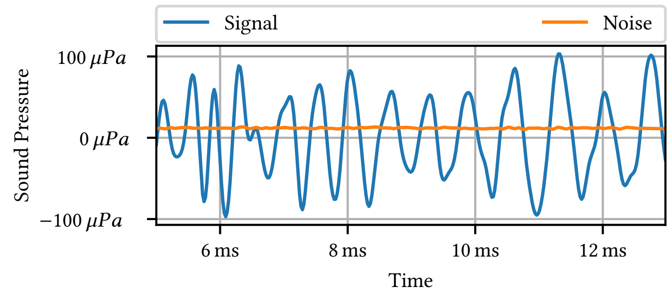
Noise reduction. During all the measurements of OAEs, the amplitude of the signal of interest is very low. Noise, either from ambient sources, or the patient (*e.g.*, due to breathing or swallowing) or from the measurement system itself, is the most limiting parameter. Due to this, all components of a OAE measurement system are carefully selected to have a low noise floor at a reasonably high sensitivity. Since smartphones are not purpose built for OAE measurement but for general audio applications, their audio behavior – *e.g.,* noise levels – is very loosely defined. However, on all platforms the signal of interest is nevertheless often too close or below the noise floor. To lower the noise floor, the measurement is repeated, while keeping the stimuli constant, and averaging the recorded signal synchronously in the time domain. For DPOAE, the averaging takes place over one FFT window size and for TEOAE buffers are chosen such that they contain one click sequence. If N is the number of recorded buffers and the noise is normally distributed and uncorrelated, the noise level is lowered by the factor $1/\sqrt{N}$ [9]. In other words, there is a 3 dB reduction in the noise level with every doubling of the number of recorded buffers.

3 SMARTPHONE-BASED SYSTEM SETUP

This section discusses the basic technical requirements for OAE-based hearing screening and how we set up our smartphones. Any



(a) Recording of the non-linear protocol click sequence, where the fourth click is inverted and has three times the amplitude. This four-click sequence is repeated multiple times, to lower the noise floor.



(b) Close-up of the averaged response window with extracted signal and noise. The high frequency components of the transient evoked OAE (TEOAE) arrive first. The response amplitude is more than 80 dB weaker compared to the stimulus.

Figure 2: Example of a TEOAE recording with the LG G5 SE smartphone.

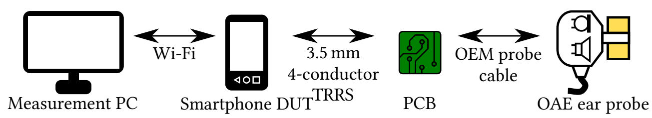


Figure 3: Measurement setup used in this paper.

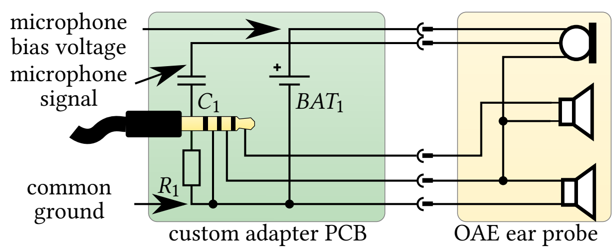


Figure 4: Schematic of the adapter PCB.

system for measuring OAEs and conducting a hearing screening test needs to at least consist of: an ear probe, one or more digital-to-analog converters (DACs) to drive the speakers in the ear probe, an analog-to-digital converter (ADC) to record the response, some kind of user interface to interact with the hearing test (*e.g.*, start/stop, display results), and a processing system to connect and drive all the components. In this paper we want to investigate whether all the components of such a system, except for the ear probe, could be replaced by a smartphone *by utilizing its headphone jack*. Additionally, we want to be able to conduct the hearing screening test without any alteration of the phone itself, be it electronically or in software, (*e.g.*, by rooting the smartphone). This allows using a wide selection of phones and would require no special skills from the user, making the solution mass deployable.

We chose to base our investigation on Android smartphones, due to their high overall market share of over 80 % [20], the availability of low-cost models and the easy access to developer resources. However, the ecosystem of Android smartphones also poses one of the main challenges: different manufacturers offer different models with many different software versions. In most cases, the manufacturer provides custom firmware and configuration for the hardware besides the Android operating system (OS). Hence, even if two smartphones share the same hardware platform and same patch level of the Android OS, they might still behave differently in terms of their audio characteristics, *e.g.*, due to different configurations of the audio amplifiers. For maximum generality, our work focuses on conducting hearing screening tests using only the APIs offered by Android to a conventional app.

3.1 Hardware and Software Setup

Figure 3 gives an overview of the hardware setup used in the experiments outlined in the following sections. A measurement personal computer (PC) controls the smartphone device under test (DUT) over a WiFi interface. In this configuration, the measurement PC runs all the processing software independently of the used smartphone. This enables us to also use other target devices, which we utilize in Section 5, to compare the smartphones with a commercially available OAE medical-grade measurement platform.

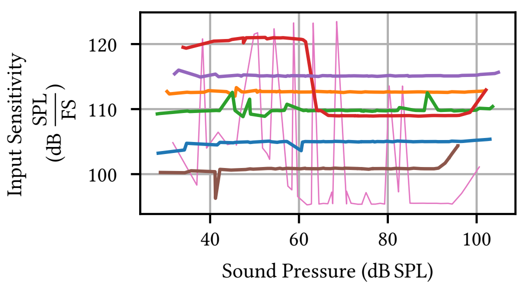
The smartphone is also connected with its 3.5 mm headphone jack to an OAE ear probe. For all measurements in this work, we used a commercially available OAE ear probe – model EP-DP produced by PATH MEDICAL GmbH [13]. To connect the OAE ear probe to the smartphone we built a custom adapter printed circuit board (PCB). Figure 4 shows the schematic of this PCB. This PCB will also handle the bias voltage for the OAE ear probe microphone. Android smartphones provide a microphone bias voltage between 1.8 V and 2.9 V [5]. However, the microphone of the used OAE ear probe requires a dedicated bias voltage and the microphone will also output a biased signal by itself. As a reliable and low-noise workaround for our experiments, we chose to put a capacitor (C_1) as a high-pass in the microphone signal path. The bias voltage for the microphone is provided by a CR2032 lithium coin cell battery. Some of the smartphones we used expected a certain direct current (DC) resistance for the probe to be correctly detected as headset, which is provided by $R_1 = 6.2\,\mathrm{k}\Omega$. The connection from our custom adapter PCB to the smartphone was provided by a regular 3.5 mm 4-conductor tip-ring-ring-sleeve (TRRS) connector cable.

3.2 Smartphone Configurations

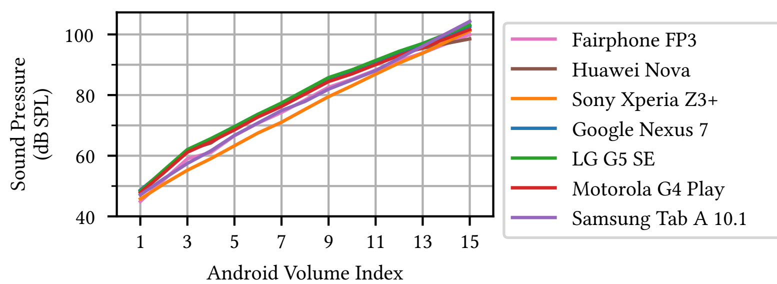
We chose to use a number of different smartphones for our experiments. Table 1 lists all the smartphones used. All the hones had the most recent system update installed and all installed apps were up-to-date. The smartphone DUTs were loaded with our custom Android app. This app offers all relevant audio functionality to conduct a hearing test via the remote control interface on the network. Using this interface, the measurement PC is used to provide the buffers to be outputted on the speakers, and to collect the corresponding buffers with the recorded signal. Further, we can control

Table 1: Overview of the smartphone DUTs used in this paper and the characterization results. The input sensitivity is measured at 1 kHz and the numeric gain error is given for a 20 dB difference in amplitude.

	Release	Android Version	Marketed as	Type	Input sensitivity Mean $dB\,\frac{SPL}{FS}$	Input sensitivity SD dB	Numeric gain error Mean dB	Numeric gain error SD dB
Fairphone FP3	Sep 2019	10	mid-range	smartphone	108.9	9.89	1.14	1.81
Huawei Nova	Oct 2016	7.0	mid-range	smartphone	100.9	0.04	0.30	0.35
Sony Xperia Z3+	May 2015	7.1.1	mid-range	smartphone	112.7	0.07	-0.17	0.35
Google Nexus 7	Jul 2013	6.0.1	mid-range	phablet	105.0	0.26	0.45	0.35
LG G5 SE	Apr 2016	7.0	mid-range	smartphone	109.8	0.25	1.69	1.22
Motorola G4 Play	May 2016	7.1.1	budget	smartphone	114.9	5.61	0.35	0.34
Samsung Tab A 10.1	May 2016	8.1.0	budget	tablet	115.1	0.04	0.23	0.33



(a) Input sensitivity dependent on the sound pressure.



(b) The resulting average sound pressure for a full scale sine wave in the ear simulator DUT dependent on the Android volume index.

Figure 5: Results of the characterization measurement in the AEC 304 ear simulator at 1 kHz.

volume levels and query other information from the smartphone. Having a non-mains connected setup for our experiments made the measurements more robust against some noise sources and also ensured the safety of the human subjects being tested.

Our app is built to use only standard calls to the Android API. According to the Android Compatibility Definition [5], only a small set of audio recording parameters must be supported by all devices, while many other settings are optional. All devices implementing audio recording, must support the "voice recognition" capture mode. In this mode the device is expected to turn off noise reduction audio processing and automatic gain control (AGC), if present. Further, some specifications regarding a flat amplitude versus frequency response, linearity of amplitude changes and total harmonic distortion (THD) are given. To be able to easily compare results across all tested devices, the sampling parameters used in the Android API was set to 16 bit mono linear pulse-code modulation (LPCM) at a sampling rate of 44,100 Hz, even though individual devices might support other modes too. For the output we used the same LPCM and sampling rate settings.

The latency in the audio stream between playback and recording is only very loosely specified. The devices must achieve an output latency (time between providing data to the Android API until measurable on headphone jack) and input latency (reverse of output latency) of maximum 500 ms each. However, the playback DACs and recording ADC are driven by the same clock source in the audio codec integrated circuit (IC) of the smartphone, so that there will be no drift between individual samples, just a fixed temporal offset once the audio streams are started. As a result, any temporal offset in the samples will only be noticeable as phase shift when doing frequency analysis, which is done in most of the measurements in this work. The only exception is the measurement of TEOAE, where some analysis and windowing was done in the time domain. However, recovering the sample shift can be achieved by simply identifying the highest amplitude transient, which indicates a specific click in the non-linear protocol.

During an OAE measurement, the sound pressure of the stimuli must be set to a certain level. This was ensured in our work using two different means – the Android volume index and numerical attenuation. The Android "volume index" is a unitless integer number. Each increment of this number increases the gain of the audio stream relatively by an unknown step. This is usually implemented on the smartphone by setting the hardware gain in the audio codec IC. Generally, the volume index should be set as low as possible, such that the output amplifiers have the smallest gain and will produce as little distortion as possible. Further, volume adjustments can be achieved by setting the numerical amplitude of the digital signal sent to the Android API. The numerical attenuation should be as low as possible (as close to full scale as possible), to have the lowest volume index setting.

In summary, there are a number of characteristics in a smartphone that we need to know prior to conducting an OAE measurement. These characteristics distinguish different phones and determine their suitability as a platform for hearing screening.

4 SMARTPHONE CHARACTERIZATION

In this section, we will identify the characteristics of the smartphone DUTs. This experiment is designed to quantify the input and output sensitivities of the audio subsystem of the smartphone DUTs and will focus on evaluating their linearity. Values given in dB FS (full scale) are referenced against the maximum digital signal value of the used LPCM encoding (0 dB FS $= 2^{15} - 1$). Sound pressure level (SPL) is a conventional unit for the pressure of sound and is referenced to approximately the threshold of human hearing at 1 KHz (0 dB SPL = 20 μPa).

4.1 Experimental Setup

The smartphone DUTs were connected to the OAE ear probe, which was inserted into a Larson Davis AEC 304 ear simulator [3]. This ear simulator resembles the acoustic properties of the human ear canal and has a microphone at the "ear drum". In this experiment a pure tone sine wave was generated, which was output by the smartphone DUTs DAC into the speakers in the ear probe. The sound pressure signal was captured by the ear simulator microphone, which was connected to a Larson Davis System 824 sound level meter (SLM) [4]. The calibrated SLM was then queried on its digital interface, to read the absolute sound pressure, the frequency and a corresponding THD value at the ear simulator "ear drum". Additionally, the smartphone DUT recorded the probe microphone, and the digital signal level was calculated.

The following parameters were tested, resulting in 300 individual measurement per DUT:

$$f = \{0.5, 1, 2, 4, 8\}\ \text{kHz},\ volume\ index = \{1 \dots 15\},$$
$$L_{num} = \{0, -20\}\ \text{dB FS},\ stereo\ channel = \{\text{L}, \text{R}\}$$

L_{num} is the numerical output level. The parameters where chosen to receive a full frequency and amplitude response in the relevant range.

4.2 Characterization Results

Input Sensitivity. In this section we evaluate the input for linearity across differing levels of sound pressure and an absolute sensitivity value for later use in the OAE experiments in Section 5.

The SLM provides absolute sound pressure values, which can be used to establish the input sensitivity of the smartphone DUT with our OAE ear probe connected. By taking the ratio of the levels of the SLM (in dB SPL) and the smartphone DUT (in dB FS), the input sensitivity can be calculated as:

$$S_{input\ smartphone} = L_{SLM} / L_{input\ smartphone}$$

This results in one sensitivity value per measurement in this experiment. Figure 5(a) shows the results, depending on the sound pressure in the ear simulator as measured by the SLM. If no AGC is active, the response should ideally be a horizontal line. This would indicate that the digital values are linear to the applied sound pressure. The response of the Motorola G4 Play was problematic, where a jump was observed, and the Fairphone FP 3 had an erratic behavior. The rising "tails" at high sound pressures, were found to be either clipping or compressed and were far beyond the sound pressures used when measuring OAE. Table 1 shows the resulting sensitivity for each smartphone DUT when only using values taken at 1 kHz and between 50 dB SPL and 80 dB SPL. These values were used as calibration offsets in the actual OAE measurements reported in Section 5.

Numeric Gain Error. Here, we evaluate whether a change in the numerical output level results in a corresponding proportional change in sound pressure output in the ear simulator. Each measurement in this experiment was done with a numerical amplitude of 0 dB FS and was repeated with -20 dB FS. Table 1 summarizes the resulting error when comparing each of these two associated measurements. Only measurements where both sound pressure readings of the SLM were between 50 dB SPL and 80 dB SPL are used for the average and standard deviation calculation. The resulting sound pressure at the ear simulator microphone should change by the same ratio as in the numerical amplitudes. It can be seen that most smartphone DUTs have a low enough error, so that changing the output amplitude can directly and proportionally be achieved with the numerical amplitude. While the LG G5 SE and the Fairphone FP 3 have a distinctly above average error, this can be compensated by using the ear probe microphone as feedback to adjust the levels.

Volume Index Gain. As discussed in Section 3, we need to set the Android volume index to as low as possible. The amount of increase in sound pressure for each step is not specified in the Android Compatibility Definition. For this reason we analyzed the relationship between sound pressure increase and the Android volume index. Figure 5(b) shows the resulting sound pressure levels read by the SLM in the ear simulator for each volume index setting at 1 kHz. The figure shows that the absolute sound pressure exerted is within a fairly narrow band of 6 dB for each volume index setting across all smartphone DUTs.

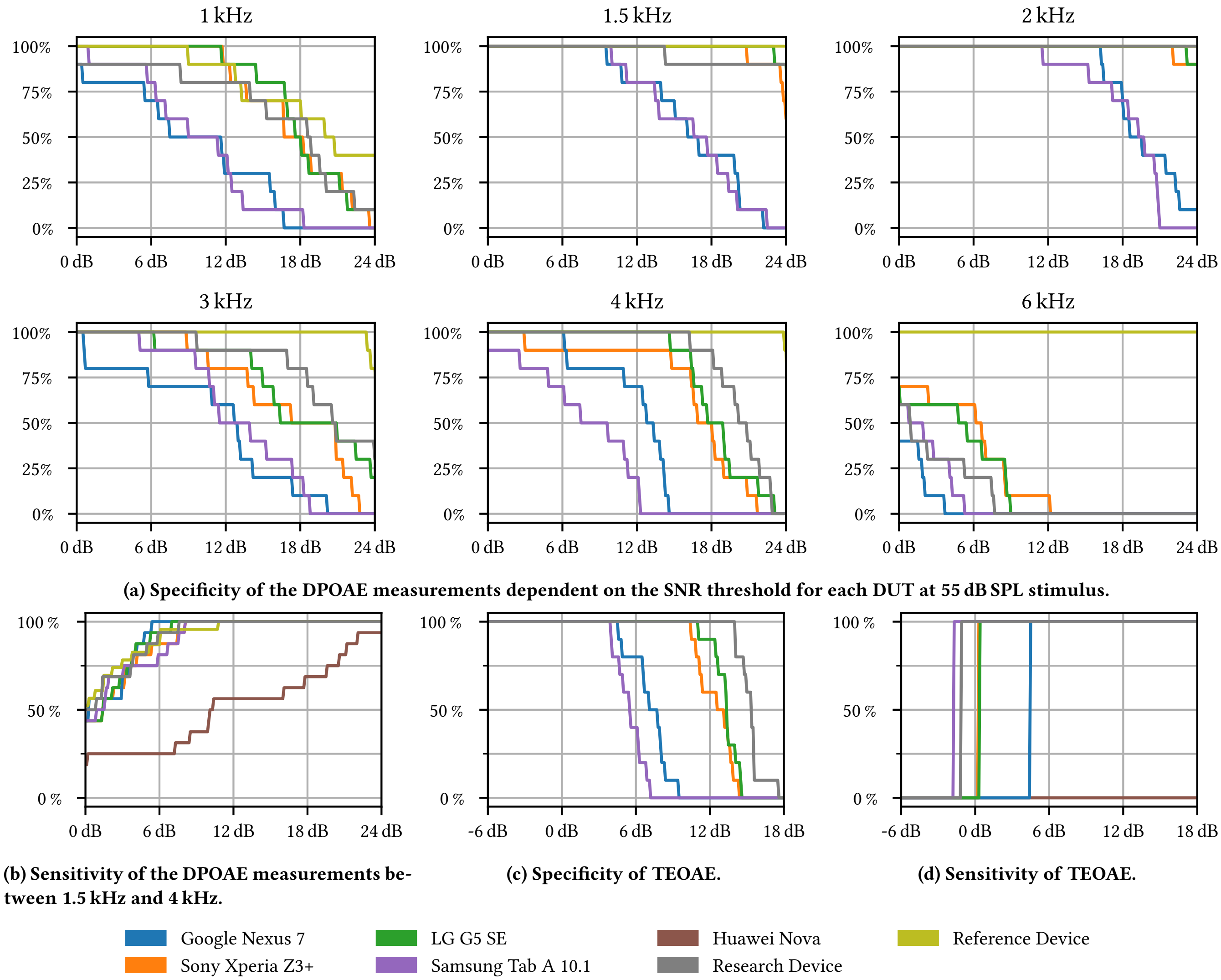
4.3 Discussion

The results of this experiment show that all smartphone DUTs are able to output the required stimuli levels in a reproducible manner. Further, we now know the "volume index" behavior for each individual device and have values for the input sensitivity at 1 kHz with an actual ear probe connected. However, noticable problems exist for some devices in the input path, which excludes them from usage in the OAE experiments. These include the Fairphone FP3 and the Motorola G4 Play.

For the actual OAE measurements, the full frequency response of the ear probe microphone was obtained in accordance with techniques outlined in Rasetshwane et al. [17]. This frequency response was then offset with each smartphone DUTs individual input sensitivity at 1 kHz. Knowing this relation is necessary to set and monitor the amplitudes of the stimuli during the OAE measurements described in the next section.

5 OAE EXPERIMENTS

In the previous section, we identified the basic characteristics of the smartphone DUT audio subsystem. However, to verify that the tested smartphone DUTs are indeed able to measure OAEs, a small study was conducted in which DPOAEs and TEOAEs were measured in human ears. Two smartphones – the Motorola G4 Play and the Fairphone FP 3 – where excluded from all following measurements, since their input behavior was unsuitable. We measured 5 adults aged 25-35, resulting in measurements in 10 ears. All measurements were preceded with an OAE measurement using



(a) Specificity of the DPOAE measurements dependent on the SNR threshold for each DUT at 55 dB SPL stimulus.

(b) Sensitivity of the DPOAE measurements between 1.5 kHz and 4 kHz.

(c) Specificity of TEOAE.

(d) Sensitivity of TEOAE.

Figure 6: Overview of the OAE measurement results. Sensitivity data was obtained in the ear simulator. The data is shown dependent on the SNR threshold. Selecting the SNR threshold for DPOAE and TEOAE is a trade off between sensitivity and specificity.

a commercially available device (PATH MEDICAL Sentiero [13]), which is referenced as the "*Reference Device*" in this section. Further a second Path Sentiero was used, referenced as the "*Research Device*", with a modified firmware so that it can be remote controlled in the same way as the smartphone DUTs.

All measurements on human ears were preceded with a measurement in the AEC 304 ear simulator. Since the ear simulator does not exhibit OAEs, it will provide data for true positive hearing loss (sensitivity). As such these measurements should result in a "refer". However, the recordings with the Huawei Nova in the ear simulator contained extremely high distortion levels in the ear simulator. So it was also not used with actual human ears. Thus, the following experiments were each repeated with only the four remaining smartphone DUTs, the research device, and the reference device.

The measurement procedure for each ear was the following. After inserting the ear probe into the ear canal, a linear chirp was

Table 2: Results of the OAE measurements with 10 normal hearing human ears. Each column shows, how many ears (out of 10) would have passed, if the given SNR threshold was applied to the measurements. For DPOAE the results are also shown if at least 3/4 or 4/4 frequencies passed.

	DPOAE				TEOAE		
	3/4 f_2		4/4 f_2				
SNR > x dB	9	12	9	12	6	9	12
Huawei Nova							
Sony Xperia Z3+	10	8	10	7	10	10	6
Google Nexus 7	10	5	8	3	8	1	0
LG G5 SE	10	9	10	9	10	10	9
Samsung Tab A 10.1	9	5	5	0	4	0	0
Research Device	10	10	10	9	10	10	10

emitted sequentially on both the ear probe speakers and the response was recorded (in-ear calibration). This frequency response was used to manually determine the quality of the probe fit and the symmetry of the speakers. For DPOAE, the frequency response was directly used to set the output gain (numeric and Android volume index), such that the primary stimuli tones were measurable at the probe microphone at a desired level. However, this direct feedback in-ear calibration may result in inadequate stimuli levels at the ear drum due to uneven sound pressure [2, 18]. For TEOAE a low amplitude click was emitted into the ear canal and the necessary gain was calculated from the response. To completely measure the health of an ear took approximately five minutes per device. All devices were measured in each ear in direct succession, if possible without removing/readjusting the ear probe.

5.1 Distortion product OAE (DPOAE)

Recordings were taken for 10 seconds at $f_2 = \{1, 1.5, 2, 3, 4, 6\}\ kHz$ and $L_2 = \{35, 45, 55, 65\}\ dB\,SPL$ resulting in 24 measurements per device per ear. The other parameters are $f_1 = f_2/1.22$ and L_1 was set with the scissor paradigm $L_1 = 0.4{\times}L_2{+}39\ dB\,SPL$ [10]. These tested frequencies and levels are common DPOAE screening parameters.

Figure 6(b) shows the sensitivity dependency of the SNR threshold for each DUTs. These measurements were done in the passive ear simulator, where no OAEs were present. Thus, all of these measurements should result in a "refer", *e.g.*, they should remain below the chosen threshold. In this data, almost no correlation to stimuli frequency or stimuli level was found. With the exception of the Huawei Nova, all DUTs performed similarly well. Figure 6(a) shows the specificity, *i.e.*, the ratio of measurements resulting in a "pass" when the hearing is actually normal, at a common screening stimuli level of $L_2 = 55$ dB SPL. Here, a connection to the stimuli parameters exists. The low specificity at 1 kHz is a limitation of DPOAEs, which are often difficult to measure at low frequencies. Further, at higher frequencies, especially at 6 kHz, the "pass" rate drops too. This is unlikely a shortcoming of the smartphone DUTs, since the research device also degrades, but not the reference device. More likely the stimuli was not at the correct level. However, results at the typical DPOAE hearing screening frequencies, between 1.5 kHz and 4 kHz were promising.

Selecting a "pass"/"refer" decision threshold on the SNR will therefore either result in better sensitivity or better specificity. Since a single measurement will only test one frequency in the cochlea, a overall "pass" is usually defined by passing at least 3/4 or even 4/4 frequencies out of $f_2 = \{1.5, 2, 3, 4\}\ kHz$. Table 2 shows the number of passing ears, when applying these criteria and also at two common SNR thresholds of 9 dB and 12 dB respectively.

5.2 Transient evoked OAEs (TEOAEs)

The stimulus for TEOAE was set to a typical level of 80 $dB\,peSPL$ and 200 non-linear click sequences were recorded per measurement, resulting in approximately 19 s of recording time for each smartphone DUT. Table 2 shows the average noise floor for each device as well as the measured SNR in the ear simulator. Again the Huawei Nova, and also to a lesser extent the Google Nexus 7, showed an increased SNR, which means that some signal was picked up, when there should have been none. Figure 6(c) shows the specificity, dependent on the SNR threshold. The reference device is not shown, since it does not output the SNR for TEOAE measurements, but all measurements resulted in a "pass" on that device.

5.3 Results and Implications

All the four smartphones used in this trial did not show an increased response in the ear simulator, neither with DPOAE nor TEOAE. This implies a good performance in terms of sensitivity. The results summarized in Table 2 show how many of the human ears could pass a screening test, when the criterion given in each column was applied. It can be seen that the performance of the LG G5 SE smartphone is almost as good as the research device.

While the performance of all smartphone DUTs was below that of the dedicated OAE screening device, the four qualified smartphones were still able to detect the the OAE signals eventually after a longer measurement duration. Further, it should be considered that our current OAE measurement software is target agnostic and was not optimized for any single device. This is in sharp contrast to the reference device, which had a target-platform specific implementation.

6 CONCLUSIONS AND FUTURE WORK

In this work we have, for the first time, demonstrated the feasibility of a smartphone-based *objective* hearing screening test. We were able to conduct our evaluations without any hardware or software modification in the phones, but by using only the standard Android API. Out of the seven smartphones we tested, we found that two (Fairphone FP 3, Motorola G4 Play) had unsuitable input behavior and one (Huawei Nova) produced unacceptable levels of distortion, making all three of them unsuitable. With the remaining four phones, we were able to conduct a OAE hearing screening trial with promising results.

Our original vision of providing potential users with an ear probe and an app so that they can user their own phone might not tenable due to the high percentage of unsuitable phones. However, providing a curated list of smartphones or providing a bundle including the smartphone and the ear probe are still viable approaches that meet our requirements, *viz.*, a mass-deployable hearing screening test that is low-cost, fast, and does not need trained medical personnel. Shipping a phone with app + headsets is currently the practice for *subjective* smartphone-based hearing tests [6]. Hence, our solution provides an *additional modality* of screening that does not require any explicit interaction with the subject. Even for providing a full solution with a smartphone and ear probe bundle, the question of manufacturing variations across two phones of the same make and model – thereby having different audio characteristics – remains open. The microphone input path characteristics are probably not closely controlled by smartphone manufacturers, to the degree they control other features. This issue needs further investigation.

Finally, this work only investigated the technical feasibility of using smartphones for objective hearing screening. To further improve usability for lay persons and *e.g.*, provide feedback on probe fit, we also plan to explore the various options that smartphones today offer in terms of connectivity and user interaction.

ACKNOWLEDGMENTS

This work was supported by Sound4All, a project funded by the Indo-German Science & Technology Centre (IGSTC).

REFERENCES

[1] Tess Bright and Danuk Pallawela. 2016. Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review. *JMIR Rehabilitation and Assistive Technologies* 3 (Dec. 2016). http://rehab.jmir.org/2016/2/e13/

[2] Karolina K. Charaziak and Christopher A. Shera. 2017. Compensating for ear-canal acoustics when measuring otoacoustic emissions. *The Journal of the Acoustical Society of America* 141, 1 (Jan. 2017), 515–531. https://doi.org/10.1121/1.4973618

[3] Larson Davis. 2021. *Model AEC304 Occluded Ear Simulator.* Retrieved June, 2021 from https://www.pcb.com/products?m=aec304

[4] Larson Davis. 2021. *System 824 Sound Level Meter.* Retrieved June, 2021 from http://www.larsondavis.com/support/slmsupport/model824

[5] Google. 2019. *Android 10 Compatibility Definition.* Retrieved June, 2021 from https://source.android.com/compatibility/10/android-10-cdd.pdf

[6] hearX Group. 2021. *Smartphone health solutions.* Retrieved June, 2021 from https://www.hearxgroup.com/

[7] D. T. Kemp. 1979. Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. *Archives of Oto-Rhino-Laryngology* 224 (March 1979), 37–45. https://doi.org/10.1007/BF00455222

[8] D. T. Kemp, P. Bray, L. Alexander, and A. M. Brown. 1986. Acoustic emission cochleography–practical aspects. *Scandinavian Audiology. Supplementum* 25 (1986), 71–95.

[9] R.S. Kulik and H. Kunov. 1995. Results of two types of averaging for distortion product otoacoustic emission measurement. In *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society.* IEEE, 979–980. https://doi.org/10.1109/IEMBS.1995.579390

[10] Peter Kummer, Thomas Janssen, Peter Hulin, and Wolfgang Arnold. 2000. Optimal L1-L2 primary tone level separation remains independent of test frequency in humans. *Hearing Research* 146 (Aug. 2000), 47–56. https://doi.org/10.1016/S0378-5955(00)00097-6

[11] Einar Laukli and Robert Burkard. 2015. Calibration/Standardization of Short-Duration Stimuli. *Seminars in Hearing* 36 (Jan. 2015), 03–10. https://doi.org/10.1055/s-0034-1396923

[12] S. B. S. Mann, S. C. Sharma, A. K. Gupta, Anu N. Nagarkar, and Dharamvir. 1998. Incidence of hearing impairment among rural and urban school going children: A survey. *Indian J Pediatr* 65 (1998), 141–145. https://doi.org/10.1007/BF02849707

[13] PATH MEDICAL. 2021. *Audiology | Diagnostic | Screening | Tracking.* Retrieved June, 2021 from https://www.pathme.de/

[14] Heidi D. Nelson, Christina Bougatsos, and Peggy Nygren. 2008. *Universal Newborn Hearing Screening: Systematic Review to Update the 2001 U.S. Preventive Services Task Force Recommendation.* Agency for Healthcare Research and Quality (US). http://www.ncbi.nlm.nih.gov/books/NBK33992/

[15] World Health Organization. 2021. *World report on hearing.* Geneva. https://apps.who.int/iris/handle/10665/339913

[16] A Ramesh, C Jagdish, M Nagapoorinima, PN Suman Rao, AG Ramakrishnan, GC Thomas, M Dominic, and A Swarnarekha. 2012. Low cost calibrated mechanical noisemaker for hearing screening of neonates in resource constrained settings. *The Indian journal of medical research* 135 (2012), 170.

[17] Daniel M. Rasetshwane and Stephen T. Neely. 2011. Calibration of otoacoustic emission probe microphones. *The Journal of the Acoustical Society of America* 130 (Oct. 2011), EL238–EL243. https://doi.org/10.1121/1.3632047

[18] Jonathan H Siegel. 2007. Calibrating otoacoustic emission probes. *Otoacoustic emissions: clinical applications* 3 (2007), 403–427.

[19] Rajesh Ranjan Singh and Rakesh Kumar Srivastava. 2019. Evaluating the Feasibility of Rolling out Universal Hearing Screening for Infants in India using Sohum, an Artificial Intelligence-driven Low Cost Innovative Diagnostic Solution. *Journal of Childhood & Developmental Disorders* 5 (2019).

[20] Statista. 2019. *Share of smartphone shipments worldwide by operating system from 2014 to 2023.* Retrieved June, 2021 from https://www.statista.com/statistics/277048/